



### **Hak cipta dan penggunaan kembali:**

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

### **Copyright and reuse:**

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Jamur merupakan salah satu tumbuhan liar yang banyak tersebar di alam liar. Lebih dari 2.000 spesies jamur ada di alam, tetapi sekitar 25 diterima secara luas sebagai makanan dan sedikit yang dibudidayakan secara komersial. Jamur telah dianggap sebagai bahan masakan *gourmet* dengan rasa yang unik. Di sisi lain, jamur juga menjadi salah satu pilihan makanan untuk bertahan hidup di alam bebas (Valverde, dkk., 2015).

Mengingat Indonesia yang memiliki iklim tropis dan mempunyai banyak hutan menjadi tempat ideal untuk tumbuhnya jamur, namun di antara jamur – jamur tersebut terdapat jamur yang berbahaya karena memiliki racun (Kementrian Lingkungan Hidup dan Kehutanan, 2018; Latha, dkk., 2018). Pengetahuan tentang jamur menjadi sangat berguna sebagai salah satu *survival kit* untuk bertahan hidup mengingat terdapat ribuan jenis jamur yang terdapat pada hutan (Stamets, 1996). Hal ini mengacu pada kejadian di tahun 2016 di California, terdapat 679 kasus keracunan jamur dan *deathcap* merupakan penyebab utamanya dikarenakan bentuknya yang cukup mirip dengan jamur umum lainnya (*California Poison Control System*, 2016).

Sejatinya jamur dapat diklasifikasikan dengan mata telanjang sesuai dengan ciri yang dimiliki oleh jamur tersebut seperti: warna insang, bentuk topi, bentuk tangkai, dan lainnya (Wibowo, 2018). Namun pengetahuan yang minim membuat orang-orang masih salah membedakan jamur yang beracun dan tidak sehingga resiko orang meninggal karena jamur masih cukup besar. Stamets (2002) menuturkan dalam bukunya bahwa jamur dapat dibedakan ke dalam 2 kelas menjadi *poisonous* dan *edible* dengan 22 atribut pembeda. Namun, terdapat kendala yang dihadapi. Jamur memiliki keberagaman, sehingga masih sulit diidentifikasi. Sebagai solusi dari permasalahan tersebut, dibutuhkan pembelajaran mesin agar klasifikasi jamur yang layak konsumsi dengan jamur yang beracun dapat dilakukan dengan *data mining*.

Menurut Witten dan Frank (2011), *data mining* dilakukan untuk mencari pola di dalam data yang nantinya dapat menghasilkan prediksi termasuk untuk klasifikasi yang dibutuhkan. *Data mining* dapat dilakukan dengan metode pembelajaran mesin. *Machine learning* atau pembelajaran mesin memungkinkan pengguna untuk memberikan ilmu kepada komputer dan membuat komputer tersebut menganalisa, dan membuat rekomendasi ataupun keputusan berdasarkan data yang diberikan sehingga merupakan pilihan tepat untuk melakukan klasifikasi (Cropper, 2019). *Machine learning* mempunyai beberapa teknik untuk melakukan pendekatan, yakni *supervised learning* dan *unsupervised learning*.

*Supervised learning* menggunakan data yang sudah diberi label dengan baik atau dengan kata lain sudah ditandai dengan jawaban yang tepat sebagai contoh pembelajaran (Iqbal, 2015). Algoritma *supervised learning* akan mempelajari data yang sudah di-*train*, dan memprediksi hasil untuk data baru. Sebagaimana jamur dapat diklasifikasikan menjadi 2 label *edible* dan *poisonous*, *supervised learning* dapat digunakan dengan baik dalam melakukan klasifikasi jamur. *Supervised learning* sendiri mempunyai beberapa algoritma, salah satunya adalah *decision tree classifier*.

Patel (2012) menegaskan bahwa *decision tree classifier* merupakan algoritma klasifikasi yang menyerupai diagram alur berstruktur pohon dengan setiap *node* yang menunjukkan *data test* pada satu atribut dan setiap cabang menunjukkan hasil dari *test* tersebut. Setiap *node* mempunyai label sesuai dengan *supervised learning*. *Decision tree* menjadi salah satu *classifier* yang baik karena langkah-langkah pembelajaran dan klasifikasi yang terbilang sederhana dan cepat. Representasi dari pembelajaran ini juga mudah dipahami oleh pengguna. Namun, *decision tree classifier* rentan terhadap kesalahan dalam masalah klasifikasi dengan banyak atribut serta *data training* dengan jumlah sedikit dan juga kasus *over-fitting* (Mitchell, 1997).

Berdasarkan masalah tersebut, *decision tree* mendapat suatu optimisasi yaitu *random forest classifier*. Menurut Starmer (2018), *Random Forest* (RF) merupakan salah satu *ensemble learning* kombinasi *decision tree* yang memberikan akurasi yang tergolong lebih tinggi karena algoritma ini akan menangani *missing values* dan tetap menjaga keakuratan data dengan jumlah besar. Breiman (2001) juga menegaskan di bukunya, bahwa kesalahan pada RF menjadi konvergen berdasarkan jumlah pohon keputusan yang terbentuk dan membuat setiap pohon tergantung pada nilai vektor yang acak diambil secara independen dan dengan distribusi yang sama untuk semua pohon keputusan. Masing–masing *decision tree* akan saling menjaga dari kesalahannya sendiri. Banyaknya *decision tree* yang terbentuk tidak akan membuat algoritma ini menjadi *over-fitting* ataupun menjadi rentan terhadap *error*.

*Random Forest* sudah diterapkan pada sejumlah penelitian seperti: deteksi kanker payudara (Aliady, dkk., 2018), penilaian mutu pendidikan (Ramadhan, 2019), analisis untuk menghasilkan informasi (Dewi, 2011), lama studi mahasiswa (Adnyana, 2016), deteksi *HIV/AIDS* (Nidhomuddin, 2015), citra alat musik tradisional (Sujaini, 2019), penilaian kredit (Adhi, 2018), pendeteksi intrusi (Teshahun, 2018), prediksi suhu perapian (Wang, 2017), dan deteksi penyakit jantung (Javeed, 2019) dengan hasil akurasi yang sangat baik dalam melakukan klasifikasi dengan akurasi berkisar antara 81.8208% sampai dengan 95.1964%.

Penelitian serupa tentang jamur telah dilakukan oleh Pranoto (Pranoto, 2019) yang melakukan klasifikasi terhadap jamur dengan menggunakan algoritma *naïve bayes*. Aplikasi yang dibuat untuk melakukan klasifikasi terhadap jamur berbasis *online* dengan tingkat akurasi dari klasifikasi yang dihasilkan sebesar 78.90730%. Sedangkan pada penelitian ini akan dilakukan proses klasifikasi jamur berdasarkan kelasnya dengan menggunakan algoritma *random forest*.

### **1.2. Rumusan Masalah**

Berdasarkan latar belakang yang telah disampaikan, maka rumusan masalah yang terjadi dalam penelitian ini adalah sebagai berikut:

- a. Bagaimana menerapkan *random forest* untuk melakukan klasifikasi agar dapat membedakan jamur yang layak konsumsi dengan jamur yang beracun?
- b. Bagaimana mengukur performa klasifikasi jamur yang menggunakan algoritma *random forest*?

### **1.3. Batasan Masalah**

Untuk menghindari perluasan masalah, maka terdapat batasan masalah dalam melakukan penelitian ini, yaitu:

- a. *Dataset* yang digunakan adalah *mushroom classification* yang diperoleh dari penelitian serupa oleh Pranoto (Pranoto, 2019) tentang prediksi kelayakan konsumsi jamur menggunakan *naïve bayes*.

- b. Jumlah sampel dari *dataset* yang digunakan sebanyak 8124 data ciri jamur dengan 23 jenis atribut, yaitu: *cap shape*, *cap surface*, *cap color*, *bruises*, *odor*, *gill attachment*, *gill spacing*, *gill size*, *gill color*, *stalk shape*, *stalk root*, *stalk surface above ring*, *stalk surface below ring*, *stalk color above ring*, *stalk color below ring*, *veil type*, *veil color*, *ring number*, *ring type*, *spore print color*, *population*, *habitat*, dan *class* sebagai label hasil dari klasifikasi jamur yang akan dilakukan.
- c. Terdapat 3916 data jamur berlabel *poisonous* (p) dan 4208 data jamur berlabel *edible* (e).

#### **1.4. Tujuan Penelitian**

Berdasarkan rumusan masalah, tujuan penelitian ini dapat dijabarkan menjadi:

- a. Untuk mengimplementasikan algoritma *random forest* untuk melakukan klasifikasi terhadap jamur agar dapat membedakan jamur yang layak konsumsi dengan jamur yang beracun pada suatu aplikasi.
- b. Mengukur performa klasifikasi jamur yang menggunakan algoritma *random forest* serta pengukuran akurasi yang dihasilkan dari algoritma tersebut.

#### **1.5. Manfaat Penelitian**

Manfaat yang dapat diambil dari penelitian ini adalah aplikasi yang dibangun diharapkan dapat membantu pengguna dalam membedakan jamur mana yang layak konsumsi dengan jamur yang beracun serta dapat meminimalisir jumlah kasus keracunan akibat konsumsi jamur liar dengan aplikasi yang diharapkan berbasis *offline* sehingga pemakaian dapat dilakukan di manapun.

## 1.6. Sistematika Penulisan

Sistematika penulisan laporan skripsi ini dapat dijabarkan dalam detail sebagai berikut:

### BAB I Pendahuluan

Pendahuluan berisi mengenai latar belakang masalah, identifikasi masalah, rumusan masalah, tujuan dan manfaat penelitian, ruang lingkup dan sistematika penulisan yang digunakan dalam penulisan laporan penelitian.

### BAB II Landasan Teori

Landasan teori berisi tentang penjelasan mengenai algoritma *Random Forest* sebagai algoritma yang digunakan untuk penelitian ini. Bab ini juga berisi landasan teori mengenai perhitungan dalam pembentukan *Decision Tree* yang digunakan pada algoritma *Random Forest*.

### BAB III Metodologi Penelitian dan Perancangan Sistem

Metodologi Penelitian dan Perancangan Sistem berisi penjelasan mengenai metodologi atau langkah-langkah dalam penyelesaian masalah, meliputi penyusunan formula dan algoritma yang akan digunakan dalam penelitian. Bab ini juga berisi perancangan *flowchart*, dan perancangan antar muka aplikasi.

### BAB IV Hasil dan Pembahasan

Implementasi dan Analisis berisi penjelasan mengenai implementasi yang sudah dibuat, dan analisis dari aplikasi yang sudah dibangun.

### BAB V Simpulan dan Saran

Berisi kesimpulan dan saran dari seluruh penelitian yang telah dilakukan yang dapat digunakan untuk pengembangan aplikasi di kemudian hari.